

Protein structure generation via folding diffusion

Year: 2022 | Citations: 249 | Authors: Kevin E. Wu, Kevin Kaichuang Yang, Rianne van den Berg, James Zou, Alex X. Lu

Abstract

The ability to computationally generate novel yet physically foldable protein structures could lead to new biological discoveries and new treatments targeting yet incurable diseases. Despite recent advances in protein structure prediction, directly generating diverse, novel protein structures from neural networks remains difficult. In this work, we present a diffusion-based generative model that generates protein backbone structures via a procedure inspired by the natural folding process. We describe a protein backbone structure as a sequence of angles capturing the relative orientation of the constituent backbone atoms, and generate structures by denoising from a random, unfolded state towards a stable folded structure. Not only does this mirror how proteins natively twist into energetically favorable conformations, the inherent shift and rotational invariance of this representation crucially alleviates the need for more complex equivariant networks. We train a denoising diffusion probabilistic model with a simple transformer backbone and demonstrate that our resulting model unconditionally generates highly realistic protein structures with complexity and structural patterns akin to those of naturally-occurring proteins. As a useful resource, we release an open-source codebase and trained models for protein structure diffusion.